SecondOpinion: Interactive Web-based Access to a Decision Model

George C. Scott, Daniel J. Cher, M.D., Leslie A. Lenert, M.D.
Divisions of General Internal Medicine and Clinical Pharmacology
Palo Alto VA Health Care System & Stanford University School of Medicine, Stanford, California

In this paper, we describe a computer architecture, which we call SecondOpinion, designed for automated, normative patient decision support over the World Wide Web. SecondOpinion custom tailors the discussion of therapy options for patients by eliciting their preferences for relevant health states via an interactive WWW interface and then integrating those results in a decision model. The SecondOpinion architecture uses a Finite State Machine representation to track the course of a patient's consultation and to choose the next action to take. The consultation has five distinct types of interactions: explanation of health states, assessment of preferences, detection and correction of errors in preference elicitations, and feedback on the implications of preference. A linear "summary model" speeds calculations of predictions from the decision model and makes it possible to dynamically calculate 95% confidence intervals for the marginal utility of each treatment option. Preferences for states are assessed in the order of their variance contribution to the models predictions in an iterative fashion. Only the states required to obtain a 95% Confidence Interval (CI) that excludes zero are assessed. In Monte Carlo simulation studies, the average number of utility assessments required for the 95% CI to exclude zero in an individual was 4.24 (SD = 1.97) out of 8 relevant heath states. The SecondOpinion architecture provides an efficient, "discussion-like" experience leading to an individual-specific treatment recommendation. It may be a cost-effective approach to bring decision analytic advice to the bedside.

INTRODUCTION AND BACKGROUND

When decision analysis was first introduced to clinical environments by Pauker and others in the late 1970s and early 1980s, it was envisioned as a bedside to aid to clinical decisions. Computers have long been thought to be instrumental in this task. Decision models are mathematical representations of problems which integrate outcome possibilities with personal utilities to predict which choice would result in the optimal cost/benefit tradeoff. It has been demonstrated that consideration of patient preferences is important in improving patients' quality of care. Further, patients appear to be interested in participating in health care decisions.

While numerous decision support systems have been developed based-on the results of decision modeling, we hypothesized that it might be possible to develop a decision support system where patients directly interacted with a decision model to obtain a treatment recommendation. By incorporating patient preference elicitation via automated computer interviewing techniques into decision models in a structured way, it should be possible to conduct a "conversation" with the patient about the implications of his or her preferences in selection of a treatment. In a previous paper⁷, we described a potential model for such decision analytic conversation between patient and computer. The computer would describe health states and elicit preferences for those states. Further, it would be able to respond in an appropriate fashion to elicited values for health states so as to reach a recommendation for each patient in the most efficient manner possible. This design would allow for the bridging of the underlying decision model with patient's values, thus leading to normative but a "humanized," custom-tailored treatment recommendation.

DESIGN CONSIDERATIONS

There have been several obstacles to the development of such a system. The current programs that use computerized interviewing methods to elicit patient preferences are structured as linear questionnaires. Not all health states will be relevant to all patients. For some patients, preferences for only one or two states may be needed to determine the preferred treatment option. Other patients, whose utilities are near the "threshold" values where the preferred treatment option changes, may need to have preference assessed for more states. Because of limits on patients attention span, it is often impractical to assess utilities for all the potentially relevant states.

Second, computation of predictions from a model needs to be rapid. If long periods of time (minutes to hours) are needed to compute recommendations, direct patient interaction will not be practical. Finally, once the recommendation is determined with the model, the computer must be able to explain the results to the patient in terms that the patient can understand. A critical aspect of this explanation is the level of

certainty. Heretofore, the "uncertainty" in a recommendation from a decision model has been estimated using one-way and two-way sensitivity analyses. The results of such analyses are qualitative and difficult to use to tailor machine generated explanations of results from the model. These obstacles and others need to be overcome to make it practical to use decision models at the bedside.

In addition to overcoming these obstacles, we had two more practical goals. First, we wanted to implement the system so that patients could access it on their own initiative. We believed the World Wide Web (WWW) would be an excellent medium to facilitate this. Second, we wanted to ensure that the system setup, use, and maintenance would not be prohibitively expensive. In face of the rapid pace with which the information base of decision systems is changing, it is imperative that such a system lend itself to equally rapid change and avoid the tremendous financial burden that conventional systems carry.⁸

IMPLEMENTATION

Discussion Model

We hypothesized that a discussion could be modeled in a data structure having a hierarchical set of five states: error, certainty, feedback, review, and assessment. Each of the states or "roles" of the system takes precedence over the states which follow it in the list. The list is cycled through until either a recommendation can be made with certainty or all preferences are assessed. Therefore, the system would first check for an error in the patient's response, such

Figure 1. Schematic of the SecondOpinion architecture.

as illogical preferences, then determine the model result's certainty (e.g. the statistical 95% confidence interval (CI) for the model). Should the degree of certainty be acceptable, the system would provide feedback and explanation about the result. Otherwise the system would provide the patient with feedback about the current result and continue to review the next topic and assess the patient's utilities for that topic.

The error state occurs whenever the patient gives an answer which is either not allowed or illogical. For instance, should the patient state that they prefer blindness to poor sight, the system would enter the error state until the error is resolved. The certainty state occurs when the patient has provided answers allow the model to determine a recommendation with an acceptable level of certainty. The feedback state occurs whenever the system is provided with new information and has reformulated its recommendation which it must then present to the patient. The review state occurs before the patient is asked to answer a question regarding their preferences. The assessment state occurs whenever the system is required to obtain information from the patient.

System Overview

Our solution was to store the user definable components of the system in a series of databases. These databases were linked with other modules to allow for manipulation of the objects and dynamic HTML generation. A schematic of the SecondOpinion system is shown in Figure 1. The discussion session is managed by a Finite State Machine (FSM) controller and a Dialog Engine.

Unlike most CGIs on the WWW which manage the client session, we used a CGI as simple messenger between the Dialog Engine and the HTTP Server. We choose to use the Dialog Engine to manage the session and generate the dynamic HTML responses.

The FSM determines the current state of the discussion based on the completeness and consistency of the patient data in the database and the uncertainty in the model parameters. We felt that a FSM model would be instrumental in allowing for the current state and the patient data to

determine what discussion paths were immediately possible. It is the job of the Dialog Engine to track where in the discussion the patient is at any time, the patient's answers, the results of the Inference Engine and to generate dynamic HTML replies. The Dialog Engine stores the patient's discussion state and information in the Patient Database. The Content Databases store methods for health state descriptions, preference assessment, inconsistency resolution, and explanation of the model results. The Dialog Engine presents all of user-defined introductory, background, and example content to the patient and then enters an iterative discussion cycle which converges on a customized recommendation. Since each iteration through the discussion cycle is driven by the patient's answers and the current discussion state, we feel that this approach should lead to customized session.

During a discussion cycle, the FSM evaluates the patient data in the database to establish consistency. If data is inconsistent, the error switch is set and the value in error identified. The Dialog Engine then generates a HTML response from user-defined segments which explains the inconsistency and prompts the patient to correct it. Once all errors are resolved, the Dialog Engine then passes the patient data to the Inference Engine.

The Inference Engine determines if the recommendation is certain based on the decision model (in which case the certainty switch in the FSM is set) and which utility contributes most to the uncertainty of the model. This utility is the next to be assessed. If the certainty state is set, then the Dialog Engine exits the discussion loop, thereby terminating the elicitation of preferences, and provides an explanation of the model results to the patient. Otherwise, the Dialog Engine continues in the cycle. If the feedback switch is set, the Dialog Engine provides an explanation about the current state of the model. Otherwise, the feedback state is skipped and

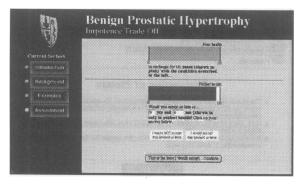


Figure 2. An example Time-tradeoff assessment screen for the health state with impotence.

the Dialog Engine enters the review state. If the review switch is set, the Dialog Engine provides user-defined review material about the utility to be assessed. After the review state, the Dialog Engine assesses the utility using a user-defined preference elicitation method (e.g. Visual Analog Scale, Time Trade-Off). An example assessment screen is shown in Figure 2.

We linked the Dialog Engine to an external Markov model that estimated a patient's quality-adjusted life expectancy. The model was implemented as an Excel (Microsoft, Redmond, WA) spreadsheet. The 95% confidence interval for predictions from the model as computed using a rapid estimation method for Markov models.7 This method uses second-order Monte Carlo simulation to compute a confidence interval for the decision model based on uncertainty in model probabilities and the population distribution of preferences for each health state. In a therapeutic conversation with a patient, the computer assesses utilities one at a time until a certain recommendation is reached and the certainty state is set. A "certain" recommendation was considered to be one which had a 95% confidence interval which excluded zero. Other degrees of certainty can be estimated directly from the probability of some other alternative actually being preferable.

Feedback about the model's current recommendation state was accomplished by providing the user with a gradient bar graph depicting the 95% CI about the recommendation certainty. A sample feedback screen is shown in Figure 3. The graph was displayed by generating dynamic HTML from user-defined segments which would resize and relocate standard graphic elements in a table.

Software Design

SecondOpinion is a suite of WWW programs that interacts with a HTTP server through Common Gateway Interface protocols. The Content Databases, DE and Patient Database where constructed in Filemaker Pro 3.0v4. The decision model was implemented as an Excel 5.0 spreadsheet and linked to the IE via AppleScript. The DE was linked to the HTTP Server with the WEB FM 3.0 CGI. WebSTAR 2.01 was used as the HTTP Server. The WWW Browser client used was Navigator 3.0. The decision to use these software packages was based on convenience and ease-of-use.

Prototype Website

We created a prototype website for providing decision analysis advice on treatment options. For the prototype, we have chosen benign prostatic hyperplasia (BPH) as the test model. We are currently assessing the architecture in WWW-based clinical studies. The URL for the SecondOpinion BPH site is preferences.stanford.edu/SecondOpinion/index.html.

VALIDATION AND EVALUATION

We tested the ability of SecondOpinion to support automated, individualized decision support in a simulation study. Monte Carlo methods were used to generate 100 patients with preferences drawn randomly from the known population distributions of the BPH model and conducted through the BPH model. Each simulated patient's discussion path was followed until a recommendation was obtained with certainty (95% CI excluded zero) or such a recommendation could not be made. The mean number of assessments was 4.24 (SD = 1.97) out of a maximum of 8. The time for SecondOpinion to calculate the 95% CI for an individual and provide feedback after a utility assessment was approximately 80 seconds.

DISCUSSION

We found that the course taken by each patient through our WWW site can be individualized using this architecture based-on patients' responses. This provides a more efficient approach to measurement of patients' preferences and, hopefully, a more conversation-like experience for the patient, leading to a better accepted and more cost-effective system. A more conventional system would assess preferences one at a time and then use sensitivity analysis to determine the optimal treatment, often without an explicit means to convey certainty. Our initial experiments demonstrate that if we "prime" the model with known population distributions for each utility and use rapid mathematical methods to determine confidence intervals, we can use these to direct the course of the interview. This allows SecondOpinion to "discuss" with the patient only issues which seem relevant for them in the decision in an interactive, conversational manner based on each utility's variance contribution to the model.

SecondOpinion, however, does not provide an individualized sequence of presentation of health states for each patient. In order to extend our approach to generate an individualized sequence, we would need to know the covariance among preferences for different states in the population. We could then re-compute the prior distribution for the utilities as yet unmeasured in the patient conditioned upon the utility values for the states already assessed. This would

allow for the most efficient and "human-like" application to patient decision support. However, it would require very large database of patient preferences to estimate conditional distributions of utilities.

The algorithm for determining the order of preference elicitation is dependent only on the variance contribution of each health state in the model and does not require any knowledge about the model itself or the clinical application area. This allows for domain independence of the architecture and implies that such systems could be developed and deployed for many different roles without changing computational mechanisms. Patients acceptance of this approach to presenting clinical materials will be explored in later experiments.

From a technical perspective, we discovered some limitations of our design. Using a FSM to represent the course of the interview created problems as far as supporting the capability for subjects to review material that had previously visited. Further, since every possible state must be enumerated beforehand and the number of states grows with every addition to the system, it quickly becomes impossible to include every possible path in the design. One solution would be to supplement the FSM with a pointer to the patient's current position.

The requirement to develop a WWW compatible system also resulted in limitation. We used a commercial CGI program to link our WWW server with a database which then coordinated other interprocess communication. We made our choice on the basis of ease of implementation and cost. However, we feel that much could be gained through the use of a more intelligent CGI which possessed higher level

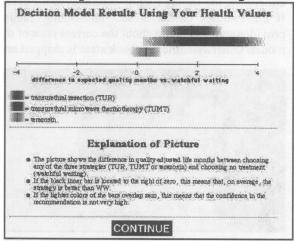


Figure 3. Example of A Feedback Screen

functionality and served as "middleware" coordinating various applications. Ideally, an "intelligent" CGI would be used to manage the entire computer-patient dialogue. One limitation is that our current CGI could put a hold-lock on the database while the decision model is being calculated in Excel. Thus for certain functions, the WWW site can only service one user at a time. This limits our capacity to service patients. A more sophisticated CGI program could prevent this problem.

As described by Slack 10, the computer is a potential tool to increase patient autonomy in medical decision making. However, we feel it is important for computerized "patient advocates" to provide more than just information. We were able to produce a system which appears to provide a customized dialogue and recommendation to patients concerning options for therapy, thereby proving that such a goal is indeed attainable. We have also implemented SecondOpinion in such a way that the patient is unaware of the mechanistic or decision analytical actions underlying the interface, yet benefits from the insight generated by a custom "discussion" driven by a decision model. This approach has allowed us to instill a "human" element into the system which we feel will become an important consideration in the future design of patient decision support systems.

FUTURE WORK

The success of this application hinges upon our ability to explain the results of decision modeling to patients. We plan to enhance the functionality of SecondOpinion with methods described by Langlotz to facilitate explanations of decision models. While our current approach for rapid approximate computation of predictions from decision models produces acceptable results, we also plan to examine other approaches such as the staged-model approach described by Rutledge. This approach might further speed computation by allowing patients to converge an approximate result with a simple decision model and continue the discussion with a more detailed model if necessary.

By providing physicians and patients with easy, costeffective access to decision models, it might be possible for them to develop more informed treatment plans and improve the quality of therapy together. Such a system could greatly contribute to increased patient autonomy in the future.

ACKNOWLEDGMENTS

This work was supported by a grant (LM 05626-02) from the National Library of Medicine, the Ambulatory Care Fellowship, Department of Veterans Affairs, and the RK Richards Medical Student Research Assistant Fund, Department of Clinical Pharmacology, Stanford University.

REFERENCES

- Pauker SG, Kassirer JP. Therapeutic decision making: a cost-benefit analysis. New Engl J Med 1975;293(5)229-34.
- ² Pauker SG, Kassirer JP. The threshold approach to clinical decision making. New Engl J Med 1980;302(20)1109-17.
- Pauker SG, Kassirer JP. Clinical decision analysis by personal computer. Arch Intern Med 1981;141(13)1831-7.
- McNeil BJ, Pauker SG, Sox HC, Tversky A. On the elicitation of preferences for alternative therapies. New Engl J Med 1982;306(21):1259-62.
- ⁵ Barry MJ, Mulley AJ, Fowler FJ, Wennberg JW. Watchful waiting vs. immediate transurethral resection for symptomatic prostatism. The importance of patient's preferences. JAMA 1988;259(20):3010-7.
- ⁶ Barry MJ, Fowler F, Jr., Mulley A, Jr., Henderson J, Jr., Wennberg JE. Patient reactions to a program designed to facilitate patient participation in treatment decisions for benign prostatic hyperplasia. Med Care 1995;33(8):771-82.
- Other DJ, Lenert LA. Rapid estimation of confidence intervals for Markov process decision models: applications in decision support systems. JAMIA (In press).
- Kasper JF, Mulley A, Jr., Wennberg JE. Developing shared decision-making programs to improve the quality of health care. QRB Qual Rev Bull 1992;18(6):183-90.
- Slack WV, Slack CW. Patient-computer dialogue. N Engl J Med 1972;286(24)1304-9.
- ¹⁰ Slack WV, Slack CW. Patient-computer dialogue. N Engl J Med 1972;286(24)1304-9.
- ¹¹ Langlotz CP, Shortliffe EH, Fagan LM. A methodology for generating computer-based explanations of decision-theoretic advice. Med Decis Making 1988;8(4):290-303.
- Rutledge, GW. VentSim: A simulation model of cardiopulmonary physiology. In: Ozbolt JG, ed. 18th Ann Symp on Comp App in Medical Care. Washington, DC: Hanley & Belfus, 1994:878-83.